THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 55, No. 4, April 1976
Printed in U.S.A.

Analysis of a Gradient Algorithm for Simultaneous Passband Equalization and Carrier Phase Recovery

By D. D. FALCONER

(Manuscript received December 11, 1975)

A two-dimensional receiver structure has been proposed, incorporating two innovations: passband equalization, which mitigates intersymbol interference, and data-directed carrier recovery and demodulation following equalization, which enables compensation of carrier frequency offset and phase jitter, but does not require transmission of a separate pilot tone with the data signal. The receiver is fully adaptive; the adjustment of the equalizer tap coefficients and of the estimate of the current channel phase shift is based on a gradient algorithm for jointly minimizing the mean squared error with respect to those parameters.

In this paper, we analyze the dynamic behavior of the deterministic gradient algorithm (where channel parameters entering into the gradient expression are assumed known in advance). The corresponding estimated gradient algorithm (where these parameters are initially unknown) has previously been studied experimentally, but is not treated here.

The first part of the present study concerns system start-up (or transient) response when the channel's phase shift is fixed. Examination of the analytical solution leads to the qualitative conclusion that, if the equalizer tap adaptation coefficient β is small relative to the phase-tracking coefficient α, the added phase estimation feature does not strongly affect the start-up behavior of the passband equalizer under typical operating conditions. Indeed, if the equalizer tap coefficients all start at zero, their evolution in the deterministic gradient algorithm is completely unaffected by the phase-tracking loop.

The second situation analyzed is the steady-state response of the system to a constant carrier frequency offset. In this case, the phase-tracking loop is found to reduce the resulting rate of rotation of the equalizer taps to about $\beta/(\alpha + \beta)$ of the original frequency offset. As a result, the degradation in system mean squared error due to frequency offset is typically quite small.

The final analysis is of the response of a linearized version of the receiver structure to sinusoidal phase jitter. When the channel's linear

distortion is not too severe and the coefficient β is small, the system mean squared error owing to phase tracking error is found to approximate that of a simple, first-order, phase-locked loop.

I. INTRODUCTION

Tbe combination of adaptive equalization and decision-directed estimation of a fixed carrier phase offset in suppressed-carrier PAM modems by means of a gradient algorithm has been suggested by Cbang[1] and by Kobayashi,[2] tbe latter also including adaptive timing recovery. Tbe receivers contemplated in those papers demodulated the received data signal prior to equalization and carrier phase estimation.

Reference 3 describes an alternative receiver configuration for two-dimensional modulated data transmission systems, combining equalization and carrier recovery. This receiver's distinction is tbat it employs a passband equalizer[4] whose reference signal consists of receiver decisions amplitude-modulating a carrier whose phase shift is the receiver's estimate of the channel phase shift. Following the passband equalizer is a demodulator which compensates for the channel's phase sbift (which may be time-varying as a result of frequency offset or phase jitter).

The receiver's estimation of the carrier phase shift is based on a decision-directed gradient algorithm for estimating a fixed phase shift, as proposed in Refs. 1, 2, 5, and 6. An advantage of the demodulator following the equalizer is that the demodulator's phase reference is delayed relative to the actual channel phase shift by only one symbol interval instead of by the entire equalizer delay as in the traditional "baseband" receiver configuration. This fact, plus the provision of a sufficiently large gain coefficient in tbe phase-tracking gradient algorithm, makes possible tracking and compensation of typical conditions of frequency offset and phase jitter that may occur on voiceband telephone channels. Computer simulations, reported in Refs. 3 and 7, have confirmed this capability.

In tbis paper, we study the dynamic behavior of tbe gradient algorithm for jointly adjusting the equalizer tap coefficients and tbe pbase estimate in each of the following situations: (*i*) start-up (transient response) for a fixed carrier phase shift; (*ii*) steady-state response to a frequency offset; (*iii*) steady-state response to sinusoidal phase jitter. Tbroughout, we consider only tbe deterministic gradient algoritbm; that is, receiver decisions are assumed perfect, and the gradient of tbe mean squared error as a function of equalizer tap coefficients and carrier reference is assumed known. A stochastic gradient algoritbm, whicb would be used in practice, bas been simulated,[3,7] but is not treated in this paper.

II. SYSTEM EQUATIONS

The transmitted two-dimensional modulated data signal is assumed to be of the form

$$s(t) = \text{Re}\left\{\sum_n A_n g(t - nT) e^{j2\pi f_c t}\right\},$$

where A_n is a two-dimensional (complex-valued) data symbol transmitted in the nth symbol interval, $g(t)$ is a band-limited baseband pulse waveform, T is the duration of a symbol interval, and f_c is the carrier frequency. The set of possible discrete complex values that each A_n can assume constitutes the signal *constellation*. Quadrature amplitude modulation (QAM) and digital phase modulation (PM) systems are familiar examples of two-dimensional modulation systems. We shall assume that successive data symbols are uncorrelated; i.e.,

$$\begin{aligned} \langle A_n A_m^* \rangle &= 1 \qquad \text{for } n = m \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

Figure 1 shows the receiver structure. The received signal, after transmission through a noisy, dispersive channel which may introduce a slowly time-varying phase shift, is passed through a phase splitter to produce parallel in-phase and quadrature components. These parallel waveforms can be represented as a single complex waveform that is sampled and passed on to a passband transversal equalizer with, say, $2N + 1$ complex-valued tap coefficients. In the nth symbol interval, when a decision is to be made on the nth data symbol, the latest $(2N + 1)$ complex-valued samples stored in the $(2N + 1)$-tap passband equalizer can be represented by the complex $(2N + 1)$-dimensional vector $\mathbf{R}_n e^{j\theta_n}$, where θ_n is the channel phase shift (assumed quasi-stationary in the nth symbol interval). A sequence $\{\theta_n\}$ changing at a constant rate with time is an example of frequency offset, while $\{\theta_n\}$ varying randomly or quasi-periodically constitutes phase jitter. Typically, the change in θ_n in one or two symbol intervals is so small as to allow us to neglect the phase-to-amplitude modulation conversion effected by filtering the sequence of incidental frequency-modulated components $\{e^{j\theta_n}\}$.

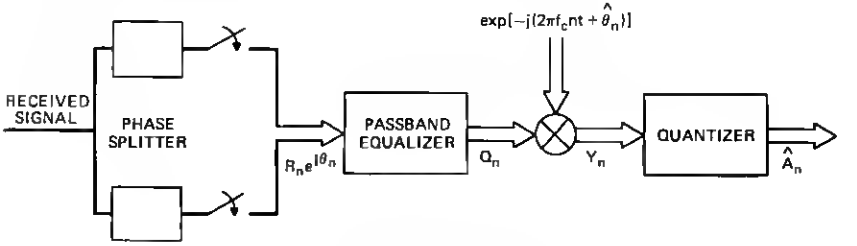

Fig. 1—Two-dimensional receiver.

The $(2N + 1)$ complex equalizer tap coefficients in the nth symbol interval are denoted by the complex $(2N + 1)$-dimensional vector $\mathbf{C}_n = (c_{-N}, \cdots, c_0, \cdots, c_N)$.* The symbol * will denote transposed complex conjugate throughout. Then the nth complex equalizer output is

$$Q_n = \mathbf{C}_n^* \mathbf{R}_n e^{j\theta_n}, \tag{1}$$

the real part heing interpreted as the in-phase component and the imaginary part as the quadrature component.

The receiver's estimate of θ_n is a real quantity denoted hy $\hat{\theta}_n$, and the demodulator output is written

$$Y_n = Q_n e^{-j(2\pi f_c nT + \hat{\theta}_n)}. \tag{2}$$

This quantity is passed into a simple quantizer to produce $\hat{A}_n$, which is the receiver's decision on A_n. Based on this decision, the complex reference signal used for updating the equalizer taps and the phase estimate is

$$\hat{Q}_n = \hat{A}_n e^{j(2\pi f_c nT + \hat{\theta}_n)}. \tag{3}$$

We define the properties of the channel in terms of expectations (denoted by $\langle\ \rangle$) with respect to the ensemhles of information symhol sequences and additive noise samples. The complex impulse response $\mathbf{X}$ is defined hy

$$\mathbf{X} = \frac{1}{\langle |A_n|^2 \rangle} \langle A_n^* \mathbf{R}_n \rangle e^{-j2\pi f_c nT}. \tag{4}$$

The positive definite Hermitian $\mathcal{A}$ matrix of the channel is defined hy

$$\mathcal{A} = \frac{1}{\langle |A_n|^2 \rangle} \langle \mathbf{R}_n \mathbf{R}_n^* \rangle. \tag{5}$$

The normalized mean squared error in the nth symbol interval is defined to be

$$\epsilon_n = \frac{1}{\langle |A_n|^2 \rangle} \langle |Q_n e^{-j(2\pi f_c nT + \hat{\theta}_n)} - A_n|^2 \rangle, \tag{6a}$$

which, hy virtue of (1), (4), and (5), can be rewritten as

$$\epsilon_n = 1 - \mathbf{X}^* \mathcal{A}^{-1} \mathbf{X} + \gamma_n, \tag{6h}$$

where $\gamma_n \equiv \mathbf{E}_n^* \mathcal{A} \mathbf{E}_n \geqq 0$ is the excess mean squared error and $\mathbf{E}_n$ is a tap-error vector,

$$\mathbf{E}_n \equiv \mathbf{C}_n e^{j(\hat{\theta}_n - \theta_n)} - \mathcal{A}^{-1} \mathbf{X}. \tag{7}$$

Since $\mathcal{A}$ is positive definite, the value of ϵ_n is a positive minimum,† $1 - \mathbf{X}^* \mathcal{A}^{-1} \mathbf{X}$, when the equalizer taps $\mathbf{C}_n$ and phase shift estimate $\hat{\theta}_n$

† The positive quantity $\mathbf{X}^* \mathcal{A}^{-1} \mathbf{X}$ is therefore less than unity, a fact which is exploited in the appendix.

are adjusted so that $\mathbf{E}_n = 0$, or

$$\mathbf{C}_n e^{j\hat{\theta}_n} = \mathcal{A}^{-1}\mathbf{X}e^{j\theta_n}. \tag{8}$$

This equation is also the condition for the gradients of ϵ_n with respect to $\mathbf{C}_n$ and $\hat{\theta}_n$ to be jointly zero; it is satisfied by an infinitude of points $(\mathbf{C}_n, \hat{\theta}_n)$.

Thus, a gradient algorithm can be used to adjust the tap coefficients $\mathbf{C}_n$ and phase estimate $\hat{\theta}_n$ recursively toward optimal values. The equations governing the evolution of $\{\mathbf{C}_n\}$ and $\{\hat{\theta}_n\}$ are[3]

$$\mathbf{C}_{n+1} = (I - \beta\mathcal{A})\mathbf{C}_n + \beta\mathbf{X}e^{-j(\hat{\theta}_n - \theta_n)} \tag{9}$$

and

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \alpha \operatorname{Im}\left[\mathbf{C}_n^*\mathbf{X}e^{-j(\hat{\theta}_n - \theta_n)}\right], \tag{10}$$

where I is the identity matrix and β and α are positive gain coefficients. These equations [or the equivalent equations (13) and (14)] form the basis for the results in this paper.

In practice, $\mathbf{X}$ and $\mathcal{A}$ would generally not be known in advance, and the following *stochastic gradient* algorithm,[3] involving the equalizer inputs $\mathbf{R}_n e^{j\theta_n}$, outputs Q_n, and modulated decisions $\hat{Q}_n$, would replace the deterministic gradient algorithm described by eqs. (9) and (10).

$$\mathbf{C}_{n+1} = \mathbf{C}_n - \beta\mathbf{R}_n e^{j\theta_n}(Q_n^* - \hat{Q}_n^*). \tag{11}$$

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \frac{\alpha}{|A_n|^2} \operatorname{Im}\,(Q_n\hat{Q}_n^*). \tag{12}$$

These are coupled stochastic difference equations, since successive vectors $\{\mathbf{R}_n\}$ are correlated random variables. Simple stochastic gradient algorithms have been studied by Widrow.[8] The application to equalizer adaptation, where no phase recovery is involved and under the assumption that the $\{\mathbf{R}_n\}$ are uncorrelated, has been studied by Ungerboeck,[9] by Gersho,[10] and by Gitlin, Mazo, and Taylor.[11] The extension to correlated vectors $\{\mathbf{R}_n\}$ has been introduced by Daniell.[12]

That the algorithm specified hy (11) and (12) converges and can perform satisfactorily is confirmed hy the computer simulations reported in Refs. 3 and 7. Analysis of the stochastic gradient algorithm is complicated by the possihility of a cycle-slipping phenomenon as in phase-lock loop systems. References 5 and 6 deal with continuous-time, decision-directed, phase-locked loops in the ahsence of adaptive equalization.

However, insight can he gained by studying instead the deterministic gradient algorithm of (9) and (10), since the estimated gradient algorithm can be interpreted as implicitly performing the averaging involved in determining $\mathbf{X}$ and $\mathcal{A}$, provided the signal-to-noise ratio is high and the gain coefficients α and β are sufficiently small.

Using definition (7), we can rewrite the coupled difference equations as

$$\mathbf{E}_{n+1} = (I - \beta\mathcal{Q})\mathbf{E}_n e^{j(\hat{\Delta}_{n+1}-\Delta_{n+1})} + \mathcal{Q}^{-1}\mathbf{X}(e^{j(\hat{\Delta}_{n+1}-\Delta_{n+1})} - 1) \quad (13)$$

and

$$\hat{\Delta}_{n+1} = \alpha \operatorname{Im}(\mathbf{E}_n^*\mathbf{X}), \quad (14)$$

where

$$\Delta_{n+1} = \theta_{n+1} - \theta_n \qquad \text{and} \qquad \hat{\Delta}_{n+1} = \hat{\theta}_{n+1} - \hat{\theta}_n.$$

III. SYSTEM START-UP WITH FIXED CHANNEL PHASE SHIFT

In this section, we study the behavior of the deterministic gradient algorithm during start-up, assuming the channel's phase shift is fixed: $\theta_n = 0$.† General theorems tell us that, if the initial error and the coefficient of the gradient algorithm are small enough, convergence is guaranteed.[13] However, we are interested in sharper results for the specific problem at hand.

The solution of (13) and (14) will depend on the initial choice of $\mathbf{E}_0$ (or $\mathbf{C}_0$) and $\hat{\theta}_0$. It is interesting to consider first the special case $\mathbf{C}_0 = 0$, the all-zero vector; i.e., $\mathbf{E}_0 = -\mathcal{Q}^{-1}X$. In this case,

$$\hat{\Delta}_1 = -\alpha \operatorname{Im}[\mathbf{X}^*\mathcal{Q}^{-1}\mathbf{X}] = 0,$$

since $\mathcal{Q}$ is Hermitian, and

$$\mathbf{E}_1 = -(I - \beta\mathcal{Q})\mathcal{Q}^{-1}\mathbf{X}.$$

Continuing, it is easy to show that

$$\Delta_n = 0 \qquad \text{for all } n$$

and that

$$\mathbf{E}_n = -(I - \beta\mathcal{Q})^n\mathcal{Q}^{-1}\mathbf{X}. \quad (15)$$

Thus, at least for this special all-zero starting condition, the estimated carrier phase shift $\hat{\theta}_n$ does not change at all and the start-up behavior of the deterministic algorithm is exactly the same as that of the passband equalizer alone.[4]

Let us now consider the more general case, when $\mathbf{E}_0$ is not necessarily equal to the right-hand side of (15) for some $n \geqq 0$. We remark that the mathematical formulation of this start-up situation will be basically the same as that of a system transient caused by an abrupt change in the channel's carrier phase shift.

Expression (6b) for the normalized mean squared error involves the positive definite quadratic form $\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n \equiv \gamma_n$. We can bound this term

† There is no loss of generality in assuming a fixed phase shift of zero, since any nonzero fixed phase-shift factor $e^{j\theta}$ can be incorporated in the complex channel impulse response $\mathbf{X}$.

and study its evolution hy writing down a recursive expression for it and upper-hounding the right-hand side of that expression. Using (13) with $\Delta_{n+1} = 0$ for all n, we can write

$$\gamma_{n+1} = \mathbf{E}_n^*(I - \beta\mathcal{A})\mathcal{A}(I - \beta\mathcal{A})\mathbf{E}_n + \mathbf{X}^*\mathcal{A}^{-1}\mathbf{X}|e^{j\hat{\Delta}_{n+1}} - 1|^2 + 2\,\mathrm{Re}\,\{\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}(1 - e^{-j\hat{\Delta}_{n+1}})\}. \quad (16)$$

The right-hand side of expression (16) is upper-bounded in the appendix. The derivation of the bound requires the following assumptions ahout the channel and algorithm parameters.

Assumption (1): The initial value $\gamma_0 \equiv \mathbf{E}_0^*\mathcal{A}\mathbf{E}_0$ is less than unity. This condition is fulfilled, for example, if $\mathbf{C}_0 = \mathbf{0}$; i.e., $\mathbf{E}_0 = -\mathcal{A}^{-1}\mathbf{X}$, for then $\gamma_0 = \mathbf{X}^*\mathcal{A}^{-1}\mathbf{X} \leqq 1$, since the positive quadratic form $\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X}$, which is one minus the minimum mean squared error, must be less than unity.

Assumption (2): $\alpha < \alpha_0$, where α_0 is the solution of

$$\alpha_0(1 + \sqrt{\gamma_0}) = 2\,\mathrm{sinc}\,(\alpha_0\sqrt{\gamma_0}),$$

where

$$\mathrm{sinc}\,\theta = \frac{\sin\theta}{\theta}.$$

Assumption (3): Let the maximum and minimum eigenvalues of the positive definite Hermitian matrix $\mathcal{A}^*$ be denoted respectively by $\lambda_{\max}$ and $\lambda_{\min}$. Then the gain coefficient β must satisfy

$$0 < \beta < \frac{2\lambda_{\min}}{\lambda_{\max}^2(1 + \epsilon_0^2)},$$

where ϵ_0^2 is defined in terms of α by

$$\alpha\left(1 + \sqrt{\gamma_0} + \frac{1}{\epsilon_0^2}\right) = 2\,\mathrm{sinc}\,(\alpha\sqrt{\gamma_0}), \qquad \alpha < \alpha_0.$$

Figure 2 illustrates the solution of the equations defining ϵ_0^2 and α_0. For example, if we assume $\alpha = 0.5$ and $\gamma_0 = 1$, then α_0 is 0.88 and ϵ_0^2 is 0.543.

The upper hound obtained in the appendix is

$$\gamma_{n+1} \equiv \mathbf{E}_{n+1}^*\mathcal{A}\mathbf{E}_{n+1} \leqq \mathbf{E}_n^*\mathcal{A}\mathbf{E}_n - 2\beta\mathbf{E}_n^*\mathcal{A}^2\mathbf{E}_n + \beta^2(1 + \epsilon_0^2)\mathbf{E}_n^*\mathcal{A}^3\mathbf{E}_n. \quad (17a)$$

An explicit bound on γ_{n+1} is obtained hy first weakening (17a) using (41) of the appendix to ohtain

$$\gamma_{n+1} \leqq (1 - 2\beta\lambda_{\min} + \beta^2(1 + \epsilon_0^2)\lambda_{\max}^2)\gamma_n, \quad (17b)$$

so that

$$\gamma_{n+1} \leqq (1 - 2\beta\lambda_{\min} + \beta^2(1 + \epsilon_0^2)\lambda_{\max}^2)^n\gamma_0. \quad (17c)$$

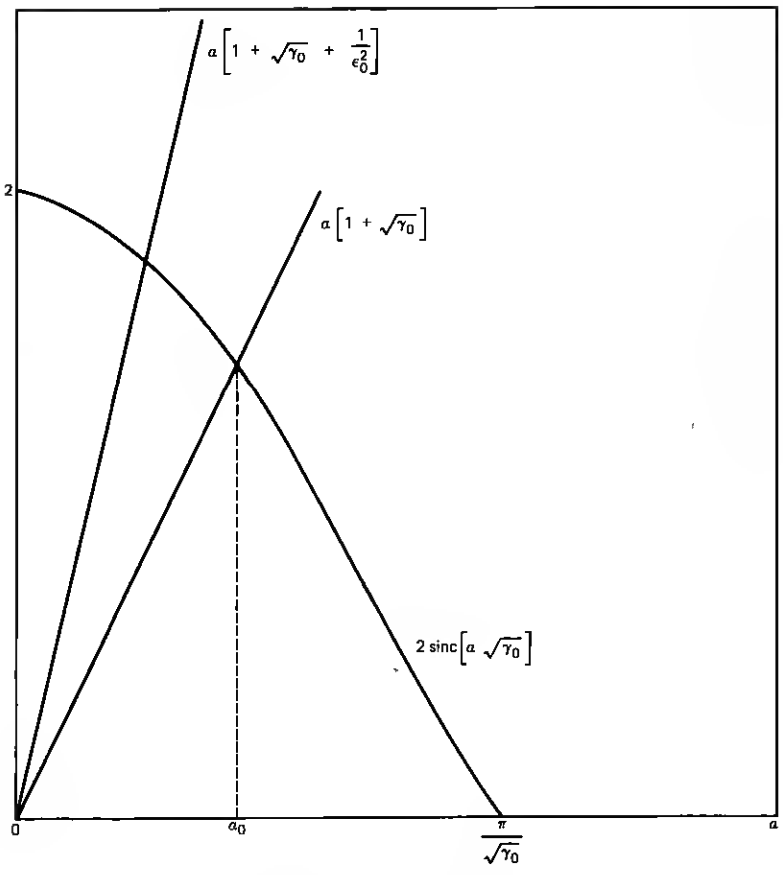

Fig. 2—Illustration of the definitions of ϵ_0^2 and a_0.

In the ahsence of phase tracking, $\alpha = \hat{\Delta}_n = 0$, and the mean squared error at step $n + 1$ of the deterministic gradient algorithm is ohtained directly from expression (15)[9,10,14] as

$$\gamma_{n+1} = \sum_i \lambda_i (1 - \beta\lambda_i)^{2n} |\mathcal{E}_{0i}|^2, \tag{18}$$

where the summation is over all the eigenvalues of the matrix $\mathfrak{A}$, the $\{\lambda_i\}$ are the set of eigenvalues, and $\mathcal{E}_{0i}$ is the inner product of $\mathbf{E}_0$ with the normalized ith eigenvector.

Comparison of the upper hound (17c) for the joint equalizing and phase-tracking receiver and the exact expression (18) for the equalizer alone yields some insight into the penalty in convergence rate imposed hy the additional phase-tracking algorithm. Consider an example where

all $\{\lambda_i\}$ (and therefore λ_{max} and λ_{min}) are equal to a common value λ. This would represent the case of a channel witb delay distortion but not amplitude distortion (flat Nyquist equivalent frequency cbaracteristic). Then inequality (17c) becomes

$$\gamma_{n+1} \leqq [(1 - \beta\lambda)^2 + \beta^2\lambda^2\epsilon_0^2]^n\gamma_0, \tag{19}$$

and, recognizing that

$$\gamma_0 \equiv \sum_i \lambda_i |\mathcal{E}_{0i}|^2,$$

we can write equality (18) for the case of no-pbase tracking as

$$\gamma_{n+1} = (1 - \beta\lambda)^{2n}\gamma_0. \tag{20}$$

In practice, the equalizer adaptation coefficient β is small ($\beta \ll 1/\lambda$), to minimize the mean squared error resulting from a practical stochastic gradient algorithm.[9] Thus the right-hand sides of (19) and (20) sbould be nearly equal, and we conclude that an ideal gradient algorithm for joint phase tracking and equalization should not converge appreciably slower than tbe equalizer adjustment algorithm alone. An exact analytical evaluation of tbe effect of phase tracking on the convergence of a practical stocbastic gradient algorithm for a severely distorted ($\lambda_{max} \gg \lambda_{min}$) channel remains elusive. However, tbe results of this section suggest that the influence of the phase-tracking parameter α in the convergence is relatively small. This conjecture is bolstered by the experimental results summarized in Figs. 3a and 3b. A 9600-b/s two-dimensional data transmission system was simulated, employing the stochastic gradient algorithm described by eqs. (11) and (12). The transmission channel, whose frequency characteristics are sbown in Fig. 3a, was regarded as severely distorted (it violates the minimum standard for private line voiceband channel data transmission). The plots of measured mean squared error versus time for $\alpha = 0$ and for $\alpha = 0.2$ shown in Fig. 3b are very similar, indicating that little penalty in convergence rate is to be ascribed to the use of joint decision-directed phase tracking.

IV. CASE OF FREQUENCY OFFSET

In this section, we study tbe behavior of tbe system in tbe presence of frequency offset by obtaining steady-state solutions to eqs. (13) and (14) when the channel phase shift increases linearly with time; i.e., $\Delta_n = 2\pi\Delta T$, where Δ is the frequency offset. In this case, eq. (13) becomes

$$\mathbf{E}_{n+1} = (I - \beta\mathcal{A})\mathbf{E}_n e^{j(\hat{\Delta}_{n+1} - 2\pi\Delta T)} + \mathcal{A}^{-1}\mathbf{X}(e^{j(\hat{\Delta}_{n+1} - 2\pi\Delta T)} - 1). \tag{21}$$

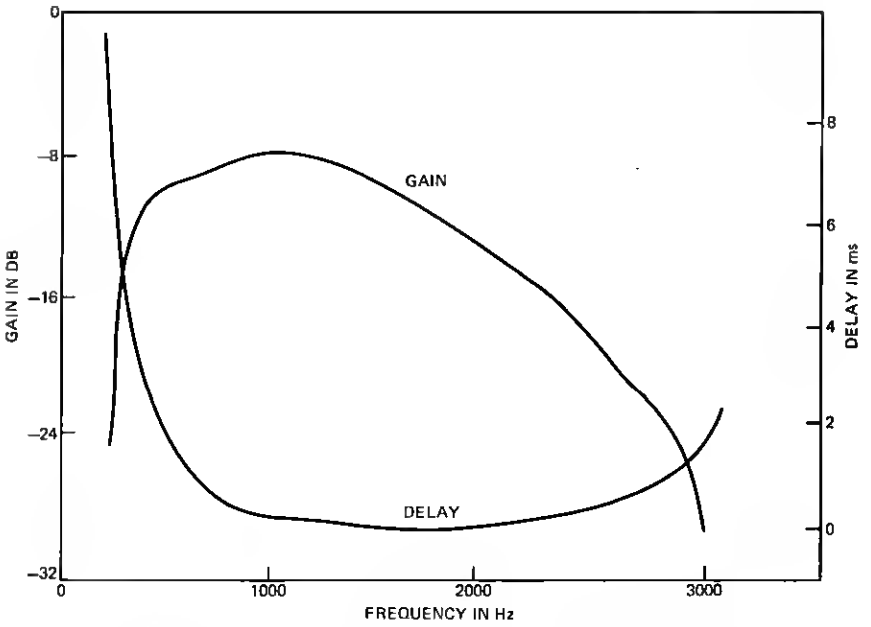

Fig. 3a—Frequency characteristics of the simulated channel.

A steady-state solution to (21) and (14) is obtained by substituting the trial solution,

$$\mathbf{E}_n = \mathbf{E}$$
$$\hat{\Delta}_n = 2\pi(\Delta + \delta)T,$$

and then solving for the fixed quantities $\mathbf{E}$ and δ. The substitution results in

$$\mathbf{E} = (e^{j2\pi\delta T} - 1)M^{-1}\mathcal{A}^{-1}\mathbf{X}, \tag{22}$$

where M is the matrix

$$M = I - e^{j2\pi\delta T}(I - \beta\mathcal{A})$$

and

$$\begin{aligned} 2\pi(\Delta + \delta)T &= \alpha \operatorname{Im}(\mathbf{E}^*\mathbf{X}). \\ &= \alpha \operatorname{Im}[(e^{-j2\pi\delta T} - 1)\mathbf{X}^*\mathcal{A}^{-1}M^{*-1}\mathbf{X}]. \end{aligned} \tag{23}$$

It is clear from the definition of M that the eigenvectors $\{\mathbf{u}_i\}_{-N}^{N}$ of $\mathcal{A}$, which form a complete orthonormal set, are also those of M. Thus, expressing the vector $\mathbf{X}$ as a linear combination of $\mathbf{u}_i$, we write

$$\mathbf{X} = \sum_{i=-N}^{N} G_i\mathbf{u}_i,$$

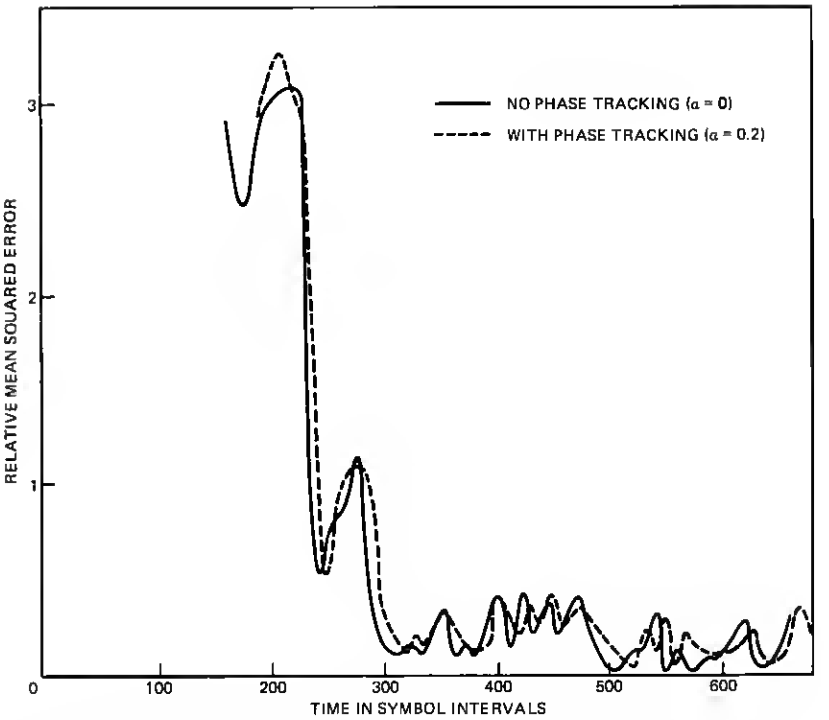

Fig. 3b—Convergence with and without phase tracking (ideal reference; all equalizer top coefficients start at zero).

we can rewrite (23) after a little algebra as

$$2\pi(\Delta + \delta)T = \alpha \operatorname{Im}\left[(e^{-j2\pi\delta T} - 1)\right] \sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i[1 - e^{-j2\pi\delta T}(1 - \beta\lambda_i)]},$$

$$= -\alpha\beta \sin 2\pi\delta T \sum_{i=-N}^{N} \frac{|G_i|^2}{1 - 2(1 - \beta\lambda_i)\cos 2\pi\delta T + (1 - \beta\lambda_i)^2}, \quad (24)$$

where $\{\lambda_i\}_{i=-N}^{N}$ are the eigenvalues of $\mathfrak{A}$ and are positive and real.

The excess mean squared error is similarly given by

$$\gamma_n = \mathbf{E}_n^* \mathfrak{A} \mathbf{E}_n = |e^{j2\pi\delta T} - 1|^2 \mathbf{X}^* \mathfrak{A}^{-1} M^* \mathfrak{A} M^{-1} \mathfrak{A}^{-1} \mathbf{X}$$

$$= 2(1 - \cos 2\pi\delta T) \sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i[1 - 2(1 - \beta\lambda_i)\cos 2\pi\delta T + (1 - \beta\lambda_i)^2]}. \quad (25)$$

Equation (24) is a transcendental equation whose solution δ is clearly not zero in general. The quantity δ may be interpreted as a bias in the receiver's estimate of the frequency offset. This "residual"

frequency offset then must be compensated for by a rotation of the equalizer complex tap coefficients at rate δ Hz.

For purposes of illustration, we again consider only a special case of a "good" channel, for which all $\lambda_i = 1$ and $\sum_i |G_i|^2 = 1$. Then (24) becomes

$$2\pi(\Delta + \delta)T = \frac{-\alpha\beta \sin 2\pi\delta T}{\beta^2 + 2(1 - \beta)(1 - \cos 2\pi\delta T)}. \tag{26}$$

Typically, $\beta \ll \alpha < 1$; for example, $\beta = 0.001$ and $\alpha = 0.2$. The left- and right-hand sides of (26) as functions of $2\pi\delta T$ are sketched in Fig. 4. Apparently in the region of intersection, $2\pi\delta T \ll \beta$ and $\sin 2\pi\delta T \approx 2\pi\delta T$. Solving (26) with this approximation yields

$$2\pi(\Delta + \delta)T \approx \frac{-2\pi\alpha}{\beta}\,\delta T.$$

Thus

$$\delta \approx \frac{-\beta\Delta}{\alpha + \beta}, \tag{27}$$

and the necessary rate of rotation of the equalizer taps has been reduced by a factor of $\beta/(\alpha + \beta)$, which is about 1/200 for a typical case, $\alpha = 0.2, \beta = 0.001$. The corresponding normalized excess mean squared error is

$$\mathbf{E}_n^* \mathfrak{A} \mathbf{E}_n \approx \frac{(2\pi\delta T)^2}{\beta^2 + (2\pi\delta T)^2} \approx \frac{(2\pi\Delta T)^2}{(\alpha + \beta)^2 + (2\pi\Delta T)^2}. \tag{28}$$

If $\Delta = 1$ Hz, $\alpha = 0.2$, $\beta = 0.001$, $T = 1/2400$ s. This amounts to about 10^{-4}.

V. STEADY-STATE SINUSOIDAL RESPONSE

The phase jitter process $\{\theta_n\}$ that occurs in telephone channels is typically quasi-periodic. It is thus of interest to determine the steady-state solution of the coupled difference equations (13) and (14) when the driving term $\{\theta_n\}$ is sinusoidal.

It is convenient at this point to rewrite eqs. (13) and (14) further in terms of eigenvalues and eigenvectors of the matrix $\mathfrak{A}$. Since $\mathfrak{A}$ is Hermitian, its eigenvalues $\{\lambda_i\}_{i=-N}^{N}$ are positive real and its eigenvectors $\{\mathbf{\mu}_i\}_{i=-N}^{N}$ form an orthonormal set which is a basis in $2N + 1$-dimensional space. Using these properties and expressing the vectors $\mathbf{E}_n$ and $\mathbf{X}$ as linear combinations of the $\{\mathbf{\mu}_i\}$,

$$\mathbf{E}_n = \sum_{i=-N}^{N} \mathcal{E}_{ni}\mathbf{\mu}_i$$

$$\mathbf{X} = \sum_{i=-N}^{N} G_i\mathbf{\mu}_i,$$

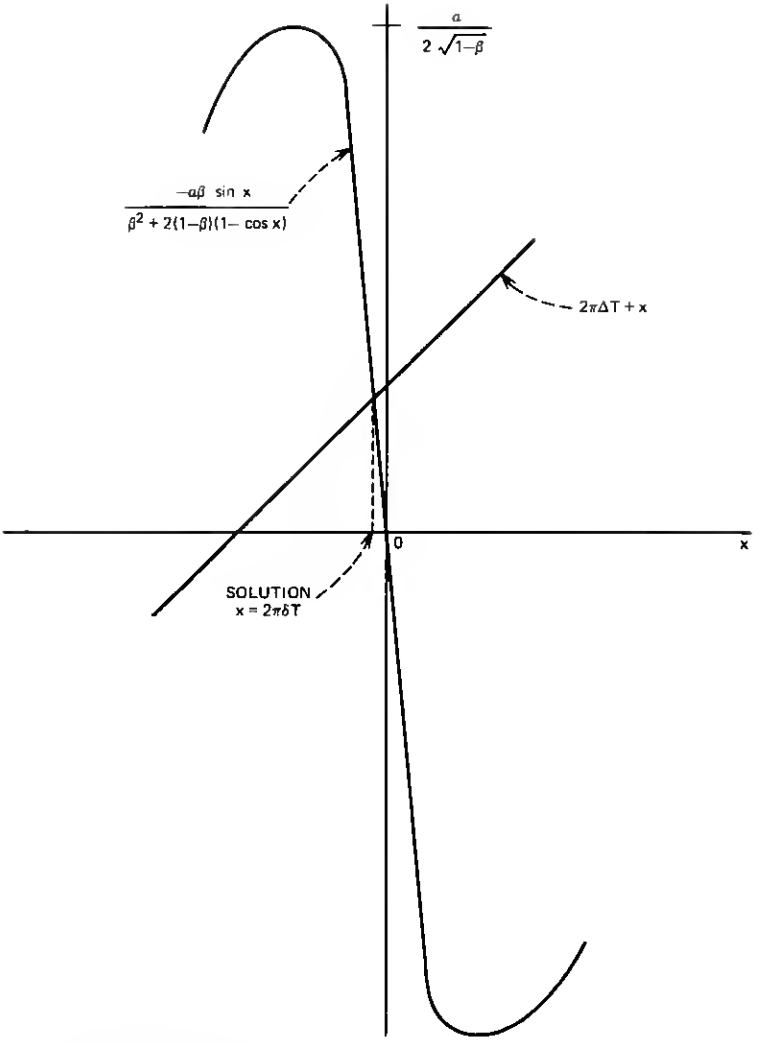

Fig. 4—Illustration of the solution of

$$2\pi(\Delta + \delta)T = \frac{-\alpha\beta \sin 2\pi\delta T}{\beta^2 + 2(1-\beta)(1-\cos 2\pi\delta T)}.$$

we can write (13) and (14) as

$$\mathcal{E}_{(n+1)i} = (1 - \beta\lambda_i)\,\mathcal{E}_{ni} e^{j(\hat{\Delta}_{n+1} - \Delta_{n+1})} + \frac{G_i}{\lambda_i}\,(e^{j(\hat{\Delta}_{n+1} - \Delta_{n+1})} - 1) \qquad -N \leqq i \leqq N \quad (29)$$

and

$$\hat{\Delta}_{n+1} = \alpha \sum_{i=-N}^{N} \operatorname{Im}\,(\mathcal{E}_{ni}^{*} G_i). \quad (30)$$

We now make the following change of variable in (29) and (30). Define

$$\mathcal{E}_{ni}^{*} G_i e^{j(\hat{\theta}_n - \theta_n)} = u_{ni} + jv_{ni}. \tag{31}$$

Then we can write the real and imaginary parts of (29) as

$$u_{(n+1)i} = (1 - \beta\lambda_i)u_{ni} + \frac{|G_i|^2}{\lambda_i}[\cos(\hat{\theta}_n - \theta_n) - \cos(\hat{\theta}_{n+1} - \theta_{n+1})], \quad -N \leqq i \leqq N \tag{32}$$

and

$$v_{(n+1)i} = (1 - \beta\lambda_i)v_{ni} + \frac{|G_i|^2}{\lambda_i}[\sin(\hat{\theta}_n - \theta_n) - \sin(\hat{\theta}_{n+1} - \theta_{n+1})] \quad -N \leqq i \leqq N, \tag{33}$$

and we can write (30) in the form

$$\hat{\theta}_{n+1} - \hat{\theta}_n = \alpha \sum_{i=-N}^{N} [v_{ni}\cos(\hat{\theta}_n - \theta_n) - u_{ni}\sin(\hat{\theta}_n - \theta_n)]. \tag{34}$$

Equations (32), (33), and (34) are a set of nonlinear coupled difference equations. In particular, eq. (34) is reminiscent of the equation governing a discrete-time, first-order, phase-locked loop. We shall solve linearized versions of (32), (33), and (34). Assuming the steady-state error angle $(\hat{\theta}_n - \theta_n)$ for $n \gg 1$ is very small, we replace $\cos(\hat{\theta}_n - \theta_n)$ by 1 and $\sin(\hat{\theta}_n - \theta_n)$ by $(\hat{\theta}_n - \theta_n)$. Then (32) becomes

$$\begin{aligned} u_{(n+1)i} &= (1 - \beta\lambda_i)u_{ni}, \\ &= (1 - \beta\lambda_i)^{n+1}u_{0i}, \qquad -N \leqq i \leqq N, \end{aligned}$$

which approaches zero in the steady state (assuming $\beta < 1/\lambda_i$ for all i). Thus in the steady state we are left with the linearized versions of (33) and (34):

$$v_{(n+1)i} = (1 - \beta\lambda_i)v_{ni} + \frac{|G_i|^2}{\lambda_i}(\theta_{n+1} - \theta_n - \hat{\theta}_{n+1} + \hat{\theta}_n) \quad -N \leqq i \leqq N \tag{35}$$

and

$$\hat{\theta}_{n+1} - \hat{\theta}_n = \alpha \sum_{i=-N}^{N} v_{ni}. \tag{36}$$

Equations (35) and (36) are linear and can be solved for a given sequence of channel phase shifts $\{\theta_n\}$. We consider the case where the phase jitter is sinusoidal with frequency ω rad/s; i.e.,

$$\theta_n = \text{Re}\,(Je^{j\omega nT}),$$

where J is a complex constant. The solution for $\{v_{ni}\}$ is also sinusoidal:

$$v_{ni} = \text{Re}\,(V_i e^{j\omega nT}). \qquad -N \leqq i \leqq N. \tag{37}$$

Substitution of this trial solution in (35) and (36) yields a value of V_i after some algebraic manipulations.

$$V_i = \frac{J(1 - e^{j\omega T})|G_i|^2}{\lambda_i(1 - \beta\lambda_i - e^{j\omega T})\left(1 - \alpha \sum_{k=-N}^{N} |G_k|^2/[(1 - \beta\lambda_k - e^{j\omega T})\lambda_k]\right)}. \tag{38}$$

It follows from the sinusoidal variation of $\{v_{ni}\}_{i=-N}^{N}$ that the error angle $\{\hat{\theta}_n - \theta_n\}$ and the equalizer tap coefficient vector $\mathbf{C}_n$ also vary sinusoidally with frequency ω in the steady state.

The excess time-averaged mean squared error can be calculated from expression (31), (37), and (38).

$$\begin{aligned}\gamma = \langle\gamma_n\rangle &= \langle\mathbf{E}_n^* \mathfrak{A} \mathbf{E}_n\rangle \\ &= \sum_{i=-N}^{N} \lambda_i \langle|\mathcal{E}_{ni}|^2\rangle \\ &= \frac{|J|^2|1 - e^{j\omega T}|^2 S_1}{2|1 - e^{j\omega T} - \alpha S_2|^2},\end{aligned} \tag{39}$$

where

$$S_1 = \sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i|1 - \beta\lambda_i/(1 - e^{j\omega T})|^2}$$

and

$$S_2 = \sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i[1 - \beta\lambda_i/(1 - e^{j\omega T})]}.$$

The total mean squared error is, from (6b),

$$\begin{aligned}\langle\epsilon_n\rangle &= 1 - \mathbf{X}^* \mathfrak{A}^{-1} \mathbf{X} + \gamma \\ &= 1 \sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i} + \frac{|J|^2|1 - e^{j\omega T}|^2 S_1}{2|1 - e^{j\omega T} - \alpha S_2|^2}.\end{aligned} \tag{40}$$

Typically, if the overall mean squared error is close to zero,

$$\sum_{i=-N}^{N} \frac{|G_i|^2}{\lambda_i} \approx 1 \qquad \text{and} \qquad \beta\lambda_i \ll |1 - e^{j\omega T}|.$$

Then the excess mean squared error in (40) is approximately

$$\frac{|J|^2|1 - e^{j\omega T}|^2}{2|1 - e^{j\omega T} - \alpha|^2}.$$

This expression corresponds to a previously derived, approximate, mean squared error due to sinusoidal jitter in the absence of noise [see eq. (39) of Ref. 3]. That equation, valid for a first-order, phase-locked loop,

was derived ignoring the coupling between eqs. (13) and (14) and assuming perfect equalization. Calculated curves of mean squared error versus α are found in Ref. 3.

VI. CONCLUSIONS

Previous studies have shown that the functions of joint passband equalization and data-directed carrier recovery in a QAM receiver can be formulated as a gradient search algorithm. If the channel parameters entering into the expression for the gradient of the mean squared error are known, it is termed a deterministic gradient algorithm. In this paper we have analyzed the start-up behavior of the deterministic gradient algorithm and also the steady-state response to frequency offset and to sinusoidal phase jitter. The more practically motivated stochastic or estimated gradient algorithm, in which the channel parameters are initially unknown, has been studied experimentally and awaits further analytical study.

It was shown that, under typical channel conditions, when the carrier phase offset is fixed, phase tracking does not greatly slow down the start-up behavior of the deterministic gradient algorithm, at least provided the equalizer adaptation coefficient β is much less than that of the phase estimator α.

The phase estimator was first proposed as an adjunct to the passband equalizer, to mitigate the effects of too-rapid tap-coefficient rotation in the presence of channel frequency offset. It has been shown that frequency offset still causes tap rotation in the equalizer-plus-phase estimator system, but that the rate of rotation is tolerable, being on the order of $1/[1 + (\alpha/\beta)]$ times the amount of frequency offset.

The steady-state response of the linearized system to sinusoidal phase jitter was obtained. When linear distortion in the channel is not severe and the coefficient β is small, the system mean squared error due to tracking error approximates that of a first-order, phase-locked loop, as was assumed in an earlier paper.

APPENDIX

We wish to upper-bound the right-hand side of (16), given assumptions (1), (2), and (3) of Section III.

$$\mathbf{E}_{n+1}^{*}\mathcal{A}\mathbf{E}_{n+1} = \mathbf{E}_{n}^{*}(I - \beta\mathcal{A})\mathcal{A}(I - \beta\mathcal{A})\mathbf{E}_{n} + \mathbf{X}^{*}\mathcal{A}^{-1}\mathbf{X}\left|e^{j\hat{\Delta}_{n+1}} - 1\right|^{2} + 2\,\mathrm{Re}\,\{\mathbf{E}_{n}^{*}(I - \beta\mathcal{A})\mathbf{X}(1 - e^{-j\hat{\Delta}_{n+1}})\}, \quad (16)$$

where $\hat{\Delta}_{n+1}$ was given by (14).

The first term on the right-hand side can be written

$$\mathbf{E}_{n}^{*}\mathcal{A}\mathbf{E}_{n} - 2\beta\mathbf{E}_{n}^{*}\mathcal{A}^{2}\mathbf{E}_{n} + \beta^{2}\mathbf{E}_{n}^{*}\mathcal{A}^{3}\mathbf{E}_{n}.$$

The matrix $\mathcal{A}$ is positive definite and Hermitian; hence,

$$-\mathbf{E}_n^*\mathcal{A}^2\mathbf{E}_n \equiv -(\mathcal{A}^{\frac{1}{2}}\mathbf{E}_n)^*\mathcal{A}(\mathcal{A}^{\frac{1}{2}}\mathbf{E}_n) \leqq -\lambda_{\min}\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n,$$

where $\lambda_{\min}$ is the minimum eigenvalue of $\mathcal{A}$. Similarly, $\mathbf{E}_n^*\mathcal{A}^3\mathbf{E}_n \leqq \lambda_{\max}^2\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n$, where $\lambda_{\max}$ is the maximum eigenvalue. Thus we note for future reference that the first term in (16) is hounded as

$$E_n^*(I - \beta\mathcal{A})\mathcal{A}(I - \beta\mathcal{A})\mathbf{E}_n \leqq (1 - 2\beta\lambda_{\min} + \beta^2\lambda_{\max})\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n. \quad (41)$$

The second term in (16) is

$$\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X}|e^{j\hat{\Delta}_{n+1}} - 1|^2 \leqq \sin^2\frac{\hat{\Delta}_{n+1}}{2},$$

since $\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X} \leqq 1$. Upper-bounding $\sin^2(\hat{\Delta}_{n+1}/2)$ by $(\hat{\Delta}_{n+1}/2)^2$ and substituting expression (14) for $\hat{\Delta}_{n+1}$, we have

$$\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X}|e^{j\hat{\Delta}_{n+1}} - 1|^2 \leqq \alpha^2[\text{Im }(\mathbf{E}_n^*\mathbf{X})]^2. \quad (42)$$

The third term in (16) can be written as the sum of three terms.

$$2\text{ Re }\{\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}(1 - e^{-j\hat{\Delta}_{n+1}})\} = 4\text{ Re }[\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}]\sin^2\frac{\hat{\Delta}_{n+1}}{2}$$
$$- 2\text{ Im }(\mathbf{E}_n^*X)\sin\hat{\Delta}_{n+1} + 2\beta\text{ Im }(\mathbf{E}_n^*\mathcal{A}\mathbf{X})\sin\hat{\Delta}_{n+1}. \quad (43)$$

As in the inequality (42), the first term in (43) is upper-bounded hy

$$\alpha^2|\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}|[\text{Im }(\mathbf{E}_n^*\mathbf{X})]^2.$$

The matrix $I - \beta\mathcal{A}$ is Hermitian; its eigenvalues are $\{1 - \beta\lambda_i\}$, where the $\{\lambda_i\}$ are the eigenvalues of $\mathcal{A}$. Let $\lambda_{\max}$ and $\lambda_{\min}$ he the maximum and minimum eigenvalues, respectively. By assumption (3), $1 - \beta\lambda_{\max} > 0$ and thus $I - \beta\mathcal{A}$ is positive definite. Therefore,

$$|\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}| = |\mathbf{E}_n^*(I - \beta\mathcal{A})^{\frac{1}{2}}\mathcal{A}^{\frac{1}{2}}\mathcal{A}^{-\frac{1}{2}}(I - \beta\mathcal{A})^{\frac{1}{2}}\mathbf{X}|$$
$$\leqq [\mathbf{E}_n^*\mathcal{A}^{\frac{1}{2}}(I - \beta\mathcal{A})\mathcal{A}^{\frac{1}{2}}\mathbf{E}_n]^{\frac{1}{2}}[\mathbf{X}^*\mathcal{A}^{-\frac{1}{2}}(I - \beta\mathcal{A})\mathcal{A}^{\frac{1}{2}}\mathbf{X}]^{\frac{1}{2}}, \quad (44)$$

where we have used Schwartz's inequality. Using the positive definiteness of $I - \beta\mathcal{A}$ and $\mathcal{A}$, we can further upper-hound the right-hand side of (44) by

$$|\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}| \leqq (1 - \beta\lambda_{\min})^2(\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n)^{\frac{1}{2}}(\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X})^{\frac{1}{2}}$$
$$\leqq (\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n)^{\frac{1}{2}}, \quad (45)$$

since the quantities $1 - \beta\lambda_{\min}$ and $\mathbf{X}^*\mathcal{A}^{-1}\mathbf{X}$ are less than unity. Thus we have upper-hounded the first term in (43) hy

$$4\text{ Re }\{\mathbf{E}_n^*(I - \beta\mathcal{A})\mathbf{X}\}\sin^2\frac{\hat{\Delta}_{n+1}}{2} \leqq \alpha^2(\mathbf{E}_n^*\mathcal{A}\mathbf{E}_n)^{\frac{1}{2}}[\text{Im }(\mathbf{E}_n^*\mathbf{X})]^2. \quad (46)$$

After substituting for $\hat{\Delta}_{n+1}$ using eq. (14), we can express the second term in (43) as

$$-2 \text{ Im } (\mathbf{E}_n^* \mathbf{X}) \sin \hat{\Delta}_{n+1} = -2\alpha[\text{Im } (\mathbf{E}_n^* \mathbf{X})]^2 \text{ sinc } [\alpha \text{ Im } (\mathbf{E}_n^* \mathbf{X})], \quad (47)$$

where

$$\text{sinc } \theta = \frac{\sin \theta}{\theta}.$$

The third term in (43) is

$$2\beta \text{ Im } (\mathbf{E}_n^* \mathfrak{A} \mathbf{X}) \sin \hat{\Delta}_{n+1},$$

which can be upper-bounded, using (14) and the inequality $|\sin \hat{\Delta}| \leqq |\hat{\Delta}|$, by

$$2\alpha\beta(|\mathbf{E}_n^* \mathfrak{A} \mathbf{X}|)[|\text{Im } (\mathbf{E}_n^* \mathbf{X})|] \leqq \epsilon^2 \beta^2 |\mathbf{E}_n^* \mathfrak{A} \mathbf{X}|^2 + \frac{\alpha^2}{\epsilon^2} [\text{Im } (\mathbf{E}_n^* \mathbf{X})]^2 \qquad \text{for any arbitrary } \epsilon,$$

where we have used the simple inequality

$$2\alpha\beta AB \leqq \epsilon^2 \beta^2 A^2 + \frac{\alpha^2}{\epsilon^2} B^2.$$

But

$$\begin{aligned} |\mathbf{E}_n^* \mathfrak{A} \mathbf{X}|^2 &= |(\mathbf{E}_n^* \mathfrak{A}^{\frac{1}{2}})(\mathfrak{A}^{-\frac{1}{2}} \mathbf{X})|^2 \\ &\leqq (\mathbf{E}_n^* \mathfrak{A}^3 \mathbf{E}_n)[\mathbf{X}^* \mathfrak{A}^{-1} \mathbf{X}] \\ &\leqq \mathbf{E}_n^* \mathfrak{A}^3 \mathbf{E}_n, \end{aligned}$$

by Schwartz's inequality and the fact that $\mathbf{X}^* \mathfrak{A}^{-1} \mathbf{X} \leqq 1$.

Thus the third term in (43) is upper-bounded by

$$\epsilon^2 \beta^2 (\mathbf{E}_n^* \mathfrak{A}^3 \mathbf{E}_n) + \frac{\alpha^2}{\epsilon^2} [\text{Im } (\mathbf{E}_n^* \mathbf{X})]^2. \quad (48)$$

Finally, substituting (42), (46), (47), and (48) into the right-hand side of (16), we have

$$\gamma_{n+1} \equiv \mathbf{E}_{n+1}^* \mathfrak{A} \mathbf{E}_{n+1} \leqq \mathbf{E}_n^* \mathfrak{A} \mathbf{E}_n - 2\beta \mathbf{E}_n^* \mathfrak{A}^2 \mathbf{E}_n + \beta^2(1 + \epsilon^2) \mathbf{E}_n^* \mathfrak{A}^3 \mathbf{E}_n + \alpha \mathfrak{R}_n [\text{Im } (\mathbf{E}_n^* \mathbf{X})]^2, \quad (49)$$

where ϵ is arbitrary and

$$\mathfrak{R}_n = \alpha \left[1 + (\mathbf{E}_n^* \mathfrak{A} \mathbf{E}_n)^{\frac{1}{2}} + \frac{1}{\epsilon^2} \right] - 2 \text{ sinc } [\alpha \text{ Im } (\mathbf{E}_n^* \mathbf{X})]. \quad (50)$$

We make the following choice of ϵ: $\epsilon = \epsilon_0$, where ϵ_0 is defined by (with $\gamma_0 = \mathbf{E}_0^* \mathfrak{A} \mathbf{E}_0$)

$$\alpha \left(1 + \sqrt{\gamma_0} + \frac{1}{\epsilon_0^2} \right) = 2 \text{ sinc } (\alpha \sqrt{\gamma_0}). \quad (51)$$

Figure 2 is a sketch of the left- and right-hand sides of eq. (51) as functions of α for various values of ϵ_0. Equation (51) has a unique solution with $0 \leqq \epsilon_0^2 < \infty$ as long as $0 \leqq \alpha < \alpha_0$, where α_0 is defined by

$$\alpha_0(1 + \sqrt{\gamma_0}) = 2 \operatorname{sinc} (\alpha_0\sqrt{\gamma_0}).$$

Note also that, by assumption (3), the coefficient $1 - 2\beta\lambda_{\min} + \beta^2\lambda_{\max}$ of $\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n$ in the bound (41) is less than 1 and hence (49) can be weakened to

$$\mathbf{E}_{n+1}^*\mathcal{Q}\mathbf{E}_{n+1} \leqq \mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n + \mathcal{R}_n[\operatorname{Im} (\mathbf{E}_n^*\mathbf{X})]^2. \tag{52}$$

Lemma: $\mathcal{R}_n$ *is negative, and hence the sequence* $\{\gamma_n \equiv E_n^*\mathcal{Q}E_n\}$ *is monotone decreasing.*

Proof: We first observe that the sinc function in (50) defining $\mathcal{R}_n$ is even, positive, and monotone decreasing provided its argument's absolute value is less than π. But its argument is

$$\alpha \operatorname{Im} (\mathbf{E}_n^*\mathbf{X}) \leqq \alpha|\mathbf{E}_n^*\mathcal{Q}^{\frac{1}{2}}\mathcal{Q}^{-\frac{1}{2}}\mathbf{X}|.$$

This can be bounded, using Schwartz's inequality, by

$$\alpha(\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n\mathbf{X}^*\mathcal{Q}^{-1}\mathbf{X})^{\frac{1}{2}} \leqq \alpha(\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n)^{\frac{1}{2}}$$

and so

$$-\operatorname{sinc} [\alpha \operatorname{Im} (\mathbf{E}_n^*\mathbf{X})] \leqq -\operatorname{sinc} [\alpha(\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n)] \qquad \text{for } \alpha(\mathbf{E}_n^*\mathcal{Q}\mathbf{E}_n)^{\frac{1}{2}} < \pi. \tag{53}$$

In particular,

$$\alpha \operatorname{Im} (\mathbf{E}_0^*\mathbf{X}) \leqq \alpha\sqrt{\gamma_0} < \pi$$

by assumption (1), and hence we can upper-bound $\mathcal{R}_0$ by

$$\mathcal{R}_0 \leqq \alpha\left(1 + \sqrt{\gamma_0} + \frac{1}{\epsilon_0^2}\right) - 2 \operatorname{sinc} (\alpha\sqrt{\gamma_0}). \tag{54}$$

According to our choice of $\epsilon = \epsilon_0$, defined by (51), the right-hand side of (54) is zero, and so $\mathcal{R}_0 \leqq 0$. It follows from (52) that $\sqrt{\gamma_1} \leqq \sqrt{\gamma_0}$, which is less than π by hypothesis. Thus $\mathcal{R}_1$ is bounded, using (53) and $\epsilon = \epsilon_0$, by $\mathcal{R}_1 \leqq \hat{\mathcal{R}}_1$, where $\hat{\mathcal{R}}_n$ is defined by

$$\hat{\mathcal{R}}_n = \alpha\left[1 + \gamma_n^{\frac{1}{2}} + \frac{1}{\epsilon_0^2}\right] - 2 \operatorname{sinc} (\alpha\gamma_n^{\frac{1}{2}}), \tag{55}$$

and $\hat{\mathcal{R}}_0 = 0$ by the definition of ϵ_0^2. Now since $\sqrt{\gamma_1} \leqq \sqrt{\gamma_0} \leqq \pi$,

$$-2 \operatorname{sinc} (\alpha\gamma_1^{\frac{1}{2}}) < -2 \operatorname{sinc} (\alpha\gamma_0^{\frac{1}{2}})$$

and so

$$R_1 \leqq \hat{R}_1 \leqq \hat{R}_0 = 0.$$

Similarly, from (52), $\gamma_2 < \gamma_1$ and by induction

$$\gamma_n \leqq \gamma_{n-1} \leqq \cdots \leqq \gamma_0$$

and all $\mathcal{R}_n \leqq 0$.

Q.E.D.

Finally, since $\mathcal{R}_n$ is negative, we obtain the following recursive upper bound from (49):

$$\gamma_{n+1} \equiv \mathbf{E}_{n+1}^* \mathcal{A} \mathbf{E}_{n+1} \leqq \mathbf{E}_n^* \mathcal{A} \mathbf{E}_n - 2\beta \mathbf{E}_n^* \mathcal{A}^2 \mathbf{E}_n + \beta^2 (1 + \epsilon_0^2) \mathbf{E}_n^* \mathcal{A}^3 \mathbf{E}_n. \quad (56)$$

REFERENCES

1. R. W. Chang, "Joint Optimization of Automatic Equalization and Carrier Acquisition for Digital Communication," B.S.T.J., *49*, No. 6 (July–August 1970), pp. 1069–1104.
2. H. Kobayashi, "Simultaneous Adaptive Estimation and Decision Algorithm for Carrier-Modulated Data Transmission Systems," IEEE Trans. Commun. Technol., *COM-19*, No. 3 (June 1971), pp. 268–280.
3. D. D. Falconer, "Jointly Adaptive Equalization and Carrier Recovery in Two-Dimensional Digital Communication Systems," B.S.T.J., *55*, No. 3 (March 1976), pp. 317–334.
4. R. D. Gitlin, E. Y. Ho, and J. E. Mazo, "Passband Equalization for Differentially Phase-Modulated Data Signals," B.S.T.J., *52*, No. 2 (February 1973), pp. 219–238.
5. W. C. Lindsey and M. K. Simon, "Carrier Synchronization and Detection of Polyphase Signals," IEEE Trans. Commun., *COM-20*, No. 6 (June 1972), pp. 441–454.
6. M. K. Simon and J. G. Smith, "Carrier Synchronization and Detection of QASK Signal Sets," IEEE Trans. Commun., *COM-22*, No. 2 (February 1974), pp. 98–106.
7. R. R. Anderson and D. D. Falconer, "Modem Evaluation on Real Channels Using Computer Simulation," National Telecommunications Conference Record, San Diego, December 1974, pp. 877–883.
8. B. Widrow, *Adaptive Filters I: Fundamentals*, TR6764-6, System Theory Laboratory, Stanford Electronics Laboratories, Stanford University, December 1966.
9. G. Ungerboeck, "Theory on the Speed of Convergence in Adaptive Equalizers for Digital Communication," IBM J. Research and Development, November 1972, pp. 546–555.
10. A. Gersho, "Adaptive Equalization of Highly Dispersive Channels for Data Transmission," B.S.T.J., *48*, No. 1 (January 1969), pp. 55–70.
11. R. D. Gitlin, J. E. Mazo, and M. G. Taylor, "On the Design of Gradient Algorithms for Digitally-Implemented Adaptive Filters," IEEE Trans. on Circuit Theory, March 1973.
12. T. P. Daniell, "Adaptive Estimation with Mutually Correlated Training Sequences," IEEE Trans. on Systems Science and Cybernetics, *SSC-6*, No. 1 (January 1970).
13. A. A. Goldstein, *Constructive Real Analysis*, New York: Harper and Row, 1967.
14. R. W. Chang, "A New Equalizer Structure for Fast Start-Up Digital Communication," B.S.T.J., *50*, No. 6 (July–August 1971), pp. 1969–2014.